

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **07325827 A**(43) Date of publication of application: **12.12.95**

(51) Int. Cl.

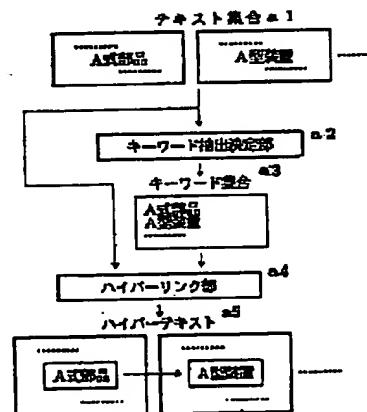
G06F 17/27**G06F 12/00****G06F 17/21****G06F 17/30**(21) Application number: **06220389**(22) Date of filing: **14.09.94**(30) Priority: **07.04.94 JP 06 69791**(71) Applicant: **mitsubishi electric corp**(72) Inventor:
**OOKASHI HITOSHI
TAKANASHI IKUKO
MINEGISHI NORIHIRO
TANAKA SATOSHI**(54) **AUTOMATIC HYPER TEXT GENERATOR**

COPYRIGHT: (C)1995,JPO

(57) Abstract:

PURPOSE: To automatically generate a hyper text even when not an index character string document but a solid text is inputted by providing a character string extraction decision part and a hyper link part.

CONSTITUTION: A keyword extraction decision part a2 first morpheme-analyzes a text set a1 and decomposes it to word units. Then, important keywords are extracted from a word set and a keyword set a3 is extracted. In this case, it is assumed that an 'A type parts' and an 'A type device' which is the synonym are extracted as the keywords. Then, the hyper link part a4 decides the keywords as hyper link origins and the keywords or the synonyms or the like of the keywords as hyper link destinations from the set a1 for the respective keywords of the set a3, generates a hyper link and outputs the hyper text a5. For instance, the hyper link for which the 'A type parts' is the hyper link origin and the 'A type device' which is the synonym of the 'A type parts' is the hyper link destination is generated.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平7-325827

(43)公開日 平成7年(1995)12月12日

(51)Int.Cl.⁸

識別記号

庁内整理番号

F I

技術表示箇所

G 0 6 F 17/27
12/00
17/21

5 4 7 H 7608-5B

9288-5L

G 0 6 F 15/ 20

5 5 0 F

9288-5L

5 7 0 D

審査請求 未請求 請求項の数16 O L (全 18 頁) 最終頁に続く

(21)出願番号 特願平6-220389

(22)出願日 平成6年(1994)9月14日

(31)優先権主張番号 特願平6-69791

(32)優先日 平6(1994)4月7日

(33)優先権主張国 日本 (J P)

(71)出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72)発明者 大塚 仁司

鎌倉市大船五丁目1番1号 三菱電機株式
会社情報システム研究所内

(72)発明者 高梨 郁子

鎌倉市大船五丁目1番1号 三菱電機株式
会社情報システム研究所内

(72)発明者 榎岸 則宏

鎌倉市大船五丁目1番1号 三菱電機株式
会社情報システム研究所内

(74)代理人 弁理士 高田 守 (外4名)

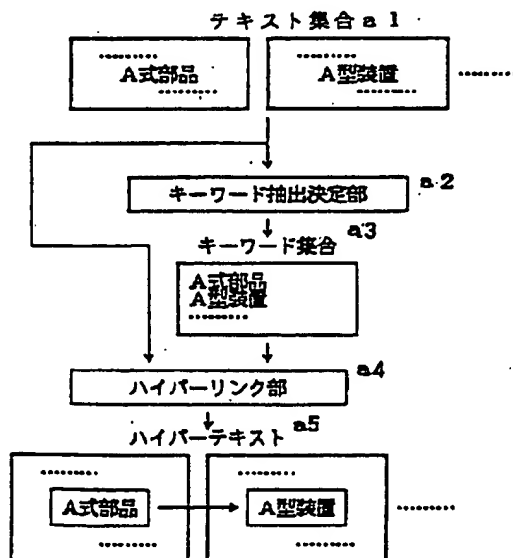
最終頁に続く

(54)【発明の名称】 ハイパーテキスト自動生成装置

(57)【要約】

【目的】 ベタのテキストからキーワードを抽出し、そのキーワードを元にハイパーリンクを自動的に設定する。

【構成】 キーワード抽出決定部 a 2 はテキスト集合 a 1 を形態素解析し、単語単位に分解してキーワード集合 a 3 を抽出する。次にハイパーリンク部 a 4 は、同一または同義のキーワードについて、テキスト集合 a 1 の中からハイパーリンク元とハイパーリンク先を決定して、ハイパーテキスト a 5 を出力する。



【特許請求の範囲】

【請求項1】 入力したテキストを所定の単位に分解し所定の文字列を抽出する文字列抽出決定部と、前記抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と一意に関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成するハイパーリンク部とを備えたことを特徴とするハイパーテキスト自動生成装置。

【請求項2】 前記ハイパーリンク部は、前記ハイパーリンク元文字列と前記ハイパーリンク先文字列を決定し両者に関連付けるハイパーリンク情報を出力するハイパーリンク決定部と、前記ハイパーリンク情報により前記ハイパーリンクを生成するハイパーリンク生成部とを備えたことを特徴とする請求項1記載のハイパーテキスト自動生成装置。

【請求項3】 前記文字列抽出決定部は、前記テキストを入力しハイパーリンク可能な文字列候補を抽出する文字列抽出部と、前記文字列候補を所定の順序に並べユーザの指定によりハイパーリンクする文字列を決定する文字列決定部とを備えたことを特徴とする請求項1記載のハイパーテキスト自動生成装置。

【請求項4】 前記ハイパーリンク決定部は、前記テキスト中の前記抽出された文字列をハイパーリンク元とし、そのテキストとは別のテキストのタイトル文字列に出現する前記ハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定することを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【請求項5】 前記ハイパーリンク決定部は、前記テキスト中の前記抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列が最も多く出現する前記ハイパーリンク元のテキストとは別のテキストのタイトル文字列をハイパーリンク先として決定することを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【請求項6】 前記ハイパーリンク決定部は、前記テキスト中の前記抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定することを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【請求項7】 前記ハイパーリンク生成部は、前記テキストのハイパーリンク元文字列とハイパーリンク先文字列を前記テキストのアドレスにより一意に決める情報及びハイパーリンク元文字列とハイパーリンク先文字列の対応関係の情報とを含んだハイパーリンク情報を記憶するハイパーリンク情報記憶部と、そのハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成するアドレスハイパーリンク生成部とを備えたことを特徴とする請求項2記載のハイパーテキス

ト自動生成装置。

【請求項8】 前記ハイパーリンク決定部は、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力し、前記ハイパーリンク生成部は、前記マークアップテキストを記憶するマークアップテキスト記憶部と、そのマークアップテキスト記憶部に記憶された前記マークアップテキストの前記タグ情報によりハイパーリンクを生成するタグハイパーリンク生成部とを備えたことを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【請求項9】 前記ハイパーリンク決定部は、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力し、前記ハイパーリンク生成部は、前記マークアップテキストを記憶するマークアップテキスト記憶部と、前記タグ情報をテーブルとして記憶するタグ情報記憶部と、そのタグ情報記憶部に記憶された前記タグ情報によりハイパーリンクを生成するタグテーブルハイパーリンク生成部とを備えたことを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【請求項10】 入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力する文書構造解析部と、前記マークアップテキストを入力し指定された所定の分割単位に前記マークアップテキストを分割して前記タグ情報を付加した部分マークアップテキスト群であるカード集合を出力するとともに、前記カード集合におけるハイパーリンク元文字列とハイパーリンク先文字列を定めたハイパーリンク情報とを出力するカード分割部と、前記ハイパーリンク情報により、前記カード集合のハイパーリンクを生成しハイパーテキストを出力するハイパーリンク生成部とを備えたことを特徴とするハイパーテキスト自動生成装置。

【請求項11】 前記ハイパーリンク生成部は、前記ハイパーリンク情報を記憶するハイパーリンク情報記憶部を備え、そのハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成することを特徴とする請求項10記載のハイパーテキスト自動生成装置。

【請求項12】 入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力する文書構造解析部と、前記マークアップテキストを入力し指定された所定の分割単位に前記マークアップテキストを分割して前記タグ情報を付加した部分マークアップテキスト群であるカード集合を出力するカード分割部と、前記カード集合を入力し前記タグ情報が付加された文字列を抽出する文字列抽出決定部と、前記抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列

10

20

30

40

50

と前記タグ情報により関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成しハイパーテキストを出力するハイパーリンク部とを備えたことを特徴とするハイパーテキスト自動生成装置。

【請求項13】 前記ハイパーリンク部は、前記カード集合を記憶部に記憶し、その記憶部に記憶されている前記カード集合の前記タグ情報によりハイパーリンクを生成することを特徴とする請求項12記載のハイパーテキスト自動生成装置。

【請求項14】 前記ハイパーリンク部は、前記カード集合を記憶部に記憶し、前記カード集合のタグ情報をテーブルとしてタグ情報記憶部に記憶し、そのタグ情報記憶部に記憶されたタグ情報によりハイパーリンクを生成することを特徴とする請求項12記載のハイパーテキスト自動生成装置。

【請求項15】 前記ハイパーリンク決定部は、リンクを生成するときの方針を示した戦略文字列を入力し、入力された戦略文字列にもとづき適切なハイパーリンクを決定する戦略選択部を備えたことを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【請求項16】 前記ハイパーリンク決定部は、前記ハイパーリンク先文字列にもとづきハイパーリンク先のテキストの表示方法の属性を決定する属性決定部を備えたことを特徴とする請求項2記載のハイパーテキスト自動生成装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、関連情報を自動的に結び付けることができるハイパーテキスト自動生成装置に関する。

【0002】

【従来の技術】ハイパーテキストは、図26に示すように、ある情報DOC10、DOC20、DOC30中に関連箇所の所在を示すハイパーリンク情報を付加することにより、複数の情報を相互に関連付けることができる。この文字列間の関連づけを行なうことをハイパーリンクを生成するという。これにより、1つの情報から関連のある情報を容易に引き出すことができる。

【0003】例えば、図26にDOC30の文字列12から、DOC10の文字列10にハイパーリンクが生成されているハイパーテキストの様子を示す。また、DOC20の文字列11から、DOC30の文字列13にハイパーリンクを生成したい場合には、まず文字列11を選択してハイパーリンク元とし、次に文字列13を選択してハイパーリンク先を設定すればよい。従来のハイパーテキストにおいて、複数のハイパーリンク元から、1つのハイパーリンク先への関連付けは行うが、1つのハイパーリンク元から複数のハイパーリンク先への関連付けは行わない。

【0004】このようにハイパーリンクが生成されてい

れば、ウィンドウに表示された情報の中のハイパーリンク設定された文字をマウスでクリックすると、ハイパーリンク先の情報が表示され、容易に関連箇所の記述や詳細内容を参照することが可能となる。

【0005】情報処理学会第47回全国大会4W-3

「文書論理構造の解析を応用したハイパーリンク自動作成支援システム」では、これらのハイパーリンク生成をすべてユーザが行なうのは大変なため、文書の構造を解析し、「章」「節」「項」のような階層構造に分けて自動的にハイパーテキストを生成することが可能である。また図や写真等の他のメディアへのハイパーリンク生成も自動的に行なっている。

【0006】特開平3-95673「文章間リンク設定方法およびその装置」においてはキーワードを抽出し、あらかじめ索引文字列を付随して持つ文章データを検索して同じ索引文字列を持つ文章データ同士のリンクを自動的に設定する。しかし入力された索引文字列文書に限られているため、例えば索引文字列のような付加的な記号のないベタのテキストをそのまま入力することはできない。また、索引文字列のついた文章同士のリンクを生成するだけなので、入力データの索引文字列の付け方によってリンク結果が違ってしまふ。

【0007】

【発明が解決しようとする課題】従来のハイパーテキストは、関連箇所のハイパーリンク設定において、ハイパーリンク元とハイパーリンク先の文字列をユーザが自分で探し、会話的にハイパーリンクを指定しなければならなかった。また、複数のハイパーリンクを設定する場合には、ユーザは前述の作業を繰り返し行なわなければならない。特に大量文書においては、ハイパーテキストの手作業による作成に手間が掛かるという問題があった。

【0008】これに対して、情報処理学会第47回全国大会4W-3「文書論理構造の解析を応用したハイパーリンク自動作成支援システム」では、ハイパーテキストを一部自動化するため、作成作業が容易になるという利点がある。しかし、ここでは、文書構造に従った階層リンクと、図や表への参照リンクを自動生成するが、文字列（キーワード）同士の自動リンク生成を行っていない。従って目次から章へ、章から節へという大まかな関連付けはできるが、それ以外で同じ意味を示す箇所同士の関連付けや、あるキーワードから他の同じ意味を示す特定のキーワードへの関連付け等は行なっていない。

【0009】また、特開平3-95673「文章間リンク設定方法およびその装置」においてもハイパーテキストの自動生成を行なうが、入力が索引文字列文書に限られているため、ベタのテキストをそのまま入力することはできない。従ってハイパーテキスト化したい文章ごとに入力データである索引文字列を付随しなければならないため、準備が大変である。また入力データの索引文字列によってリンクされる文章が決まるため、索引文字列

が少なくとリンクが少なくなる、という不具合が起こる。

【0010】本発明は、上記のような問題点を解消するためになされたもので、テキストをベタのまま入力して文章構造に従ってカードに分割し、キーワードを抽出し、そのキーワードを元にハイパーリンクを自動的に設定する、ハイパーテキストの自動生成を目的としている。また、ハイパーテキストの自動生成においても、文章の階層化だけでなく、キーワードからキーワードへのハイパーリンク生成を行なうことを目的としている。

【0011】

【課題を解決するための手段】第1の発明は、入力したテキストを所定の単位に分解し所定の文字列を抽出する文字列抽出決定部と、抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と一意に関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成するハイパーリンク部とを備えたものである。

【0012】第2の発明によるハイパーリンク部は、ハイパーリンク元文字列とハイパーリンク先文字列を決定し両者を関連付けるハイパーリンク情報を出力するハイパーリンク決定部と、そのハイパーリンク情報によりハイパーリンクを生成するハイパーリンク生成部とを備えたものである。

【0013】第3の発明による文字列抽出決定部は、テキストを入力しハイパーリンク可能な文字列候補を抽出する文字列抽出部と、文字列候補を所定の順序に並べユーザの指定によりハイパーリンクする文字列を決定する文字列決定部とを備えたものである。

【0014】第4の発明は、テキスト中の抽出された文字列をハイパーリンク元とし、そのテキストとは別のテキストのタイトル文字列に出現するハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定するハイパーリンク決定部を備えたものである。

【0015】第5の発明は、テキスト中の抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列が最も多く出現するハイパーリンク元のテキストとは別のテキストのタイトル文字列をハイパーリンク先として決定するハイパーリンク決定部を備えたものである。

【0016】第6の発明は、テキスト中の抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定するハイパーリンク決定部を備えたものである。

【0017】第7の発明によるハイパーリンク生成部は、テキストのハイパーリンク元文字列とハイパーリンク先文字列をテキストのアドレスにより一意に決める情報及びハイパーリンク元文字列とハイパーリンク先文字

列の対応関係の情報とを含んだハイパーリンク情報を記憶するハイパーリンク情報記憶部と、そのハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成するアドレスハイパーリンク生成部とを備えたものである。

【0018】第8の発明は、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力するハイパーリンク決定部と、マークアップテキストを記憶するマークアップテキスト記憶部と、そのマークアップテキスト記憶部に記憶されたマークアップテキストのタグ情報によりハイパーリンクを生成するタグハイパーリンク生成部とを備えたものである。

【0019】第9の発明は、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力するハイパーリンク決定部と、マークアップテキストを記憶するマークアップテキスト記憶部と、タグ情報をテーブルとして記憶するタグ情報記憶部と、そのタグ情報記憶部に記憶されたタグ情報によりハイパーリンクを生成するタグテーブルハイパーリンク生成部とを備えたものである。

【0020】第10の発明は、入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力する文書構造解析部と、マークアップテキストを入力し指定された所定の分割単位にマークアップテキストを分割してタグ情報を付加した部分マークアップテキスト群であるカード集合を出力するとともに、前記カード集合におけるハイパーリンク元文字列とハイパーリンク先文字列を定めたハイパーリンク情報とを出力するカード分割部と、ハイパーリンク情報により、カード集合のハイパーリンクを生成しハイパーテキストを出力するハイパーリンク生成部とを備えたものである。

【0021】第11の発明は、第10の発明におけるハイパーリンク生成部がハイパーリンク情報を記憶し、その記憶されたハイパーリンク情報によりハイパーリンクを生成するものである。

【0022】第12の発明は、入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力する文書構造解析部と、マークアップテキストを入力し指定された所定の分割単位にマークアップテキストを分割してタグ情報を付加した部分マークアップテキスト群であるカード集合を出力するカード分割部と、カード集合を入力しタグ情報が付加された文字列を抽出する文字列抽出決定部と、抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列とタグ情報により関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成しハイパーテキストを出力

10

20

30

40

50

するハイパーリンク部とを備えたものである。

【0023】第13の発明は、第12の発明におけるハイパーリンク部がカード集合を記憶し、記憶されているカード集合のタグ情報によりハイパーリンクを生成するものである。

【0024】第14の発明は、第12の発明におけるハイパーリンク部がカード集合のタグ情報をテーブルとして記憶し、記憶されたそのタグ情報によりハイパーリンクを生成するものである。

【0025】第15の発明によるハイパーリンク決定部は、リンクを生成するときの方針を示した戦略文字列を入力し、入力された戦略文字列にもとづき適切なハイパーリンクを決定する戦略選択部を備えたものである。

【0026】第16の発明によるハイパーリンク決定部は、ハイパーリンク先文字列にもとづきハイパーリンク先のテキストの表示方法を決定する属性決定部を備えたものである。

【0027】

【作用】第1の発明においては、文字列抽出部が入力したテキストを所定の単位に分解し所定の文字列を抽出し、ハイパーリンク部が抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と一意に関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成する。

【0028】第2の発明においては、ハイパーリンク決定部がハイパーリンク元文字列とハイパーリンク先文字列を決定し両者を関連付けるハイパーリンク情報を出力し、ハイパーリンク生成部がハイパーリンク情報によりハイパーリンクを生成する。

【0029】第3の発明においては、文字列抽出部がテキストを入力しハイパーリンク可能な文字列候補を抽出し、文字列決定部が文字列候補を所定の順序に並べユーザの指定によりハイパーリンクする文字列を決定する。

【0030】第4の発明においては、ハイパーリンク決定部が、テキスト中の抽出された文字列をハイパーリンク元とし、そのテキストとは別のテキストのタイトル文字列に出現するハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定する。

【0031】第5の発明においては、ハイパーリンク決定部が、テキスト中の抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列が最も多く出現するハイパーリンク元のテキストとは別のテキストのタイトル文字列をハイパーリンク先として決定する。

【0032】第6の発明においては、ハイパーリンク決定部が、テキスト中の抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定する。

【0033】第7の発明のハイパーリンク生成部においては、ハイパーリンク情報記憶部がテキストのハイパーリンク元文字列とハイパーリンク先文字列をテキストのアドレスにより一意に決める情報及びハイパーリンク元文字列とハイパーリンク先文字列の対応関係の情報とを含んだハイパーリンク情報を記憶し、アドレスハイパーリンク生成部がそのハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成する。

【0034】第8の発明においては、ハイパーリンク決定部が、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力し、タグハイパーリンク生成部が、マークアップテキスト記憶部に記憶されたマークアップテキストのタグ情報によりハイパーリンクを生成する。

【0035】第9の発明においては、ハイパーリンク決定部が、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力し、タグテーブルハイパーリンク生成部が、タグ情報記憶部に記憶されたタグ情報によりハイパーリンクを生成する。

【0036】第10の発明においては、文書構造解析部が、入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力し、カード分割部が、マークアップテキストを入力し指定された所定の分割単位にマークアップテキストを分割してタグ情報を付加した部分マークアップテキスト群であるカード集合を出力するとともに、カード集合におけるハイパーリンク元文字列とハイパーリンク先文字列を定めたハイパーリンク情報とを出力し、ハイパーリンク生成部がハイパーリンク情報により、カード集合のハイパーリンクを生成しハイパーテキストを出力する。

【0037】第11の発明においては、ハイパーリンク生成部が、ハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成する。

【0038】第12の発明においては、文書構造解析部が、入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力し、カード分割部が、マークアップテキストを入力し指定された所定の分割単位にマークアップテキストを分割してタグ情報を付加した部分マークアップテキスト群であるカード集合を出力し、文字列抽出決定部が、カード集合を入力しタグ情報が付加された文字列を抽出し、ハイパーリンク部が抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列とタグ情報により関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成しハイパーテキストを出力する。

【0039】第13の発明においては、ハイパーリンク部が記憶部に記憶されているカード集合のタグ情報によりハイパーリンクを生成する。

【0040】第14の発明においては、ハイパーリンク部が、カード集合のタグ情報をテーブルとしてタグ情報記憶部に記憶し、そのタグ情報記憶部に記憶されたタグ情報によりハイパーリンクを生成する。

【0041】第15の発明のハイパーリンク決定部においては、リンクを生成するときの方針を示した戦略文字列を入力し、戦略選択部が入力された戦略文字列にもとづき適切なハイパーリンクを決定する。

【0042】第16の発明のハイパーリンク決定部においては、属性決定部がハイパーリンク先文字列にもとづきハイパーリンク先のテキストの表示方法を決定する。

【0043】

【実施例】以下、本発明における実施例を図面を用いて説明する。

実施例1. 図1は、第1の実施例に係るハイパーテキスト自動生成装置を表す図である。図のa1は入力されるテキスト集合、a2はテキスト集合からキーワードを抽出するキーワード抽出決定部、a3は抽出されたキーワード集合、a4はテキスト集合に対してキーワード集合中の各キーワードに関するハイパーリンクを生成するハイパーリンク部、a5は生成されたハイパーテキストである。

【0044】次に動作について説明する。キーワード抽出決定部a2はテキスト集合a1をまず形態素解析し、単語単位に分解する。次に単語集合の中から重要なキーワードを抜き出してキーワード集合a3を抽出する。ここで、キーワードとして「A式部品」とその同義語である「A型装置」が抽出されたとする。次にハイパーリンク部a4がキーワード集合a3の各キーワードについてテキスト集合a1中からそのキーワードをハイパーリンク元とし、そのキーワード若しくはそのキーワードの同義語またはそのキーワード若しくは同義語が多く出現するテキストのタイトルなどをハイパーリンク先として決定してハイパーリンクを生成し、ハイパーテキストa5を出力する。例えば、「A式部品」がハイパーリンク元となり、「A式部品」の同義語である「A型装置」がハイパーリンク先となったハイパーリンクを生成する。

【0045】特開平3-95673「文章間リンク設定方法およびその装置」などの従来の装置においては入力索引文字列付き文書に限られており、ベタのテキストをそのまま入力することはできないが、本発明の装置ではベタのテキストをそのまま入力することができる。また、特開平3-95673の図1のリンク設定部はキーワードから文書データを指すリンクを設定するが、本発明ではキーワードからキーワードを指すリンクを設定することができる。

【0046】また、特開平3-95673「文章間リン

ク設定方法およびその装置」は文字列対テキストの自動ハイパーリンクであるが、本発明は、文字列から文字列への自動ハイパーリンクを可能にし、テキスト単位ではなく文字列単位の詳細なハイパーリンクを可能にする。

【0047】実施例2. 図2は、第2の実施例に係るハイパーテキスト自動生成装置を表す図である。図2において、a1は入力するテキスト集合、a2はキーワード抽出決定部、a3は抽出されたキーワード集合、a4はハイパーリンク部、a5は生成されたハイパーテキストである。また図2において、b1はテキスト中のハイパーリンク元キーワードとハイパーリンク先キーワードを決定しハイパーリンク情報b2を出力するハイパーリンク決定部、b3はハイパーリンク決定部b1で出力されるハイパーリンク情報b2に従ってハイパーリンクを生成するハイパーリンク生成部で、ともに図1のハイパーリンク部a4に相当するものの中に構成されている。ここで、ハイパーリンク情報b2は、一般的にハイパーリンク元文字列をハイパーリンク先文字列へ関連付けするための情報で、それぞれの文字列が一意に定義付けられ、関連付けのための対応付けがされているものである。

【0048】次に動作について説明する。テキスト集合a1からキーワード抽出決定部a2により抽出されたキーワード集合a3は、ハイパーリンク決定部b1に入力される。ハイパーリンク決定部b1では、テキスト集合a1に対し、入力されたキーワードをもとにハイパーリンク元とハイパーリンク先を決定し、ハイパーリンク情報b2を出力する。ハイパーリンク生成部b3では、ハイパーリンク情報b2に従ってハイパーリンクを生成し、ハイパーテキストa5を出力する。

【0049】このように本実施例のハイパーテキスト自動生成装置は、ハイパーリンクを決定する部分とハイパーリンクを生成する部分を分離することによって、ハイパーリンク決定とハイパーリンク生成を別のマシンで行なうことができるため、C/S構成において、サーバでハイパーリンクを決定し、クライアントでのハイパーリンク生成することが可能となる。また、ハイパーリンク決定をするアプリケーションとハイパーリンクを生成するアプリケーションを分けることができる。これはテキストに索引文字列を付加する等の前処理等を必要としないため、容易にハイパーテキスト化を行なうことができる。

【0050】また、本実施例のハイパーテキスト自動生成装置は、ハイパーリンク情報b2を一時テーブル等に蓄えるため、決定したハイパーリンク情報b2を一部変更する際には、蓄えたハイパーリンク情報b2自体を変更することによって、最初からすべての設定をやり直す手間が省ける。更に、蓄えたハイパーリンク情報b2を用いて、同じ内容のテキストに対して何度でも変更が可能である。これはテキストに前処理等を必要としない

め、容易にハイパーテキスト化を行なうことができる。

【0051】実施例3. 図3は、第3の実施例に係るハイパーテキスト自動生成装置を表す図である。この実施例は、実施例1や実施例2に入力するテキスト集合から、マークアップテキストの形にしたものを生成するもので、図3において、a1は入力するテキスト、c2はISO 8879/1986規格およびJIS X 4151/1992規格であるSGML (Standard Generalized Markup Language) における文書型定義であり、テキストを解析しマークアップするために短縮参照の機能を使う。c3は短縮参照機能をサポートしたSGMLパーサーなどの文書構造解析部であり、テキストをフルマークアップすることができる。c4はフルマークアップされたマークアップテキスト、c5は章や節などカード分割する単位を指定する分割単位指定、c6はテキストをカードに分割するカード分割部、c7は出力されるカード集合、c8はタイトル文字列に関するハイパーリンク元とハイパーリンク先を定めたタイトルハイパーリンク情報である。

【0052】次に動作について説明する。テキストa1と文書型定義c2とが文書構造解析部c3に入力されると、文書構造解析部c3はマークアップテキストc4を出力する。マークアップテキストc4は、SGML規格に則ってフルマークアップされる。たとえば、図中の「<章>」はSGML規格で開始タグと呼ばれるものであり章の論理構造の開始を表しており、「</章>」はSGML規格で終了タグと呼ばれるものであり章の論理構造の終了を表している。出力されたマークアップテキストc4と分割単位指定c5とがカード分割部c6に入力されると、カード分割部c6は開始タグおよび終了タグを手掛かりにマークアップテキストc4を部分マークアップテキスト群であるカード集合c7に分割する。

【0053】カード分割部c6の動作で特徴的なのは、分割単位指定c5で指定した章や節などの分割単位を表す章題や節題を一段上位のカードにも残すことである。たとえば節のカードに節題を含ませるのは当然として、一段上位の章のカードにも節題を残す。これによって、上位のカードから一段下位のカードへタイトル文字列を使ってハイパーリンクを自動生成することができる。自動生成のために、タイトル文字列に関するハイパーリンク元とハイパーリンク先を定めたタイトルハイパーリンク情報c8を出力する。

【0054】以上のように本実施例では、ハイパーテキストの表示単位であるカードをテキストから自動生成することができる。

【0055】実施例4. 図4は、第4の実施例に係るハイパーテキスト自動生成装置を表す図である。図において、a1は入力するテキスト、d2はテキスト中からハイパーリンク可能な文字列候補を抽出するキーワード抽

出部 d3は抽出されたハイパーリンク可能文字列候補、d4は当該文字列候補を元にユーザが実際にハイパーリンクするキーワード集合a3を決定するキーワード決定部である。ここで、キーワード抽出部d2とキーワード決定部d4は、図1及び図2におけるキーワード抽出決定部a2に相当するものの中に構成されている。

【0056】次に動作について説明する。テキストa1がキーワード抽出部d2に入力されるとキーワード抽出部d2は形態素解析および経験則によるキーワード絞り込みなどの日本語処理を行って、ハイパーリンク可能文字列候補d3を出力する。このハイパーリンク可能文字列候補d3は機械処理によって抽出された文字列なので、真にユーザがハイパーリンクしたい文字列に対して過不足および誤りがある。そこで、ハイパーリンク可能文字列候補d3をキーワード決定部d4へ入力し、ユーザが削除および追加を行って、キーワード集合a3として出力する。キーワード決定部d4は、キーワード集合a3を単にリストアップするのではなく、出現頻度やキーワードの品詞などを使ってユーザが選ぶ可能性の高いキーワードの順に提示する。

【0057】従来はハイパーリンクする文字列は人間がテキスト中から選んでいたが、本実施例では以上のように、キーワード候補がテキストから抽出され提示されるので、その中からハイパーリンクするキーワードを選べばよくなり、人間の作業が軽減される。

【0058】実施例5. 図5は、第5の実施例に係るハイパーテキスト自動生成装置を表す図である。図において、a1はテキスト集合、a3はハイパーリンクする文字列が何個か記述されたキーワード集合、b1はハイパーリンク決定部、c4はマークアップテキスト、b2はハイパーリンク元とハイパーリンク先を定めたハイパーリンク情報である。本実施例に関わるハイパーリンク先の決定アルゴリズムを図12～図17に示す。図12ではタイトル中の同一文字列へ、図13ではタイトル中の同義文字列へ、図14では同一文字列が最も多く出現するカードのタイトルへ、図15では同義文字列が最も多く出現するカードのタイトルへ、図16では同一文字列へ、図17では同義文字列へ、各々ハイパーリンクするものである。ここでハイパーリンク決定部b1は、図2におけるハイパーリンク決定部b1に相当するものである。

【0059】次に動作について説明する。テキスト集合a1とキーワード集合a3がハイパーリンク決定部b1に入力されると、各ハイパーリンク文字列についてそのハイパーリンク元文字列とハイパーリンク先文字列が各々以下の手法に従って決定され、マークアップテキストc4およびハイパーリンク情報b2として出力される。マークアップテキストc4またはハイパーリンク情報b2は、その片方だけでもハイパーリンク元およびハイパーリンク先の決定情報を含むので、ハイパーリンク決定

部b1の出力としてはマークアップテキストc4またはハイパーリンク情報b2のいずれか片方でも良い。また、図5に図示していないが、このマークアップテキストc4やハイパーリンク情報b2は、図2に示すハイパーリンク生成部b3に入力され、ハイパーリンクを生成する。

【0060】次に図12から図17に係るハイパーテキスト自動生成装置におけるハイパーリンク情報決定手法について述べる。

【0061】図12では、テキストDOC1中の「A式部品」をハイパーリンク元とし、別テキストのタイトル文字列中に「A式部品」が出現するかどうかを探し、別テキストDOC2のタイトル文字列に出現する同じ「A式部品」をハイパーリンク先としてハイパーリンクすることを決定する。すなわち、入力されたキーワード集合a3の文字列の内、いずれかのテキストのタイトル文字列中に出現する文字列がハイパーリンク元およびハイパーリンク先となったハイパーリンク情報b2が出力される。入力されたキーワード集合a3の文字列の内、いずれのテキストのタイトルにも出現しない文字列は、ハイパーリンク元にもハイパーリンク先にもならない。タイトル文字列に出現するような文字列を当該タイトル文字列へハイパーリンクすることを決定するというこの手法は、ハイパーリンク元文字列に関する詳しい情報がハイパーリンク先のテキスト中に記述されている可能性が高いと考えられ、有効な自動ハイパーリンク手法の一つである。

【0062】図12が同じ文字列どうしをハイパーリンクするのに対して、図13では「A式部品」と同じ意味の「A型装置」を探してハイパーリンク先として決定することにより、より有効なハイパーリンクを設定するものである。

【0063】図14では、テキストDOC1中の「A式部品」をハイパーリンク元とし、当該文字列と同じ「A式部品」が最も多く出現する別テキストDOC2を探して、DOC2のタイトル文字列をハイパーリンク先としてハイパーリンクすることを決定する。当該文字列と同じ「A式部品」が最も多く出現する別テキストDOC2は、ハイパーリンク元「A式部品」に関する詳しい情報が記述されているテキストである可能性が高いと考えられ、有効な自動ハイパーリンク手法の一つである。

【0064】図14では、タイトル文字列に「第・・・章」というタイトル記号を含めなかったが、これは含めても良い。

【0065】図14が同じ文字列で出現頻度を判断するのに対して、図15では「A式部品」と同じ意味の「A型装置」が最も多く出現する別テキストDOC2を探して、DOC2のタイトル文字列をハイパーリンク先として決定することにより、より有効なハイパーリンクを設定するものである。

【0066】図15では、タイトル文字列に「第・・・章」というタイトル記号を含めなかったが、これは含めても良い。

【0067】図16では、テキストDOC1中の「A式部品」をハイパーリンク元とし、当該文字列と同じ「A式部品」が出現する同一テキストDOC1または別テキストDOC2を探して、当該「A式部品」をハイパーリンク先としてハイパーリンクすることを決定する。そのようなハイパーリンク先「A式部品」として複数の候補が存在する場合は考えられるが、同一テキストDOC1の場合はハイパーリンク元文字列1の直後に現れる「A式部品」に、別テキストDOC2の場合は最も最初に出現する当該「A式部品」にハイパーリンクするという決定法を取れば良い。図16の決定方法は、同じキーワードに関する記述を順次閲覧したい場合に有効である。

【0068】図16が同じ文字列どうしをハイパーリンクするのに対して、図17では「A式部品」と同じ意味の「A型装置」を探してハイパーリンク先として決定することにより、より有効なハイパーリンクを設定するものである。

【0069】図12から図17に係るハイパーリンク先のテキスト決定手法において、そのようなテキストとして複数の候補が存在する場合は考えられる。この場合の決定法としては、ハイパーリンク元文字列が属するテキストに対してテキストの木構造階層関係上で最も近いテキストを選択するという手法を取れば良い。または、木構造階層関係上で最も上位のテキストを選択するという手法も考えられる。さらに、そのようなテキストが前方向と後方向に候補として2テキスト存在する可能性が考えられるが、前あるいは後ろのどちらかに決定するようにハイパーリンク決定部を作成すれば良い。

【0070】なお、図13、図15、図17の発明に係るハイパーテキスト自動生成装置では、ある2つの文字列が同じ意味の文字列であることを判断するために、図5のハイパーリンク決定部b1は同義語辞書またはソーラスを持つ。

【0071】実施例6、図6は、第6の実施例に係るハイパーテキスト自動生成装置を表す図である。本実施例におけるハイパーテキスト自動生成装置は、ハイパーリンク情報b2として例えば図8のようなハイパーリンクテーブルを記憶するハイパーリンク情報記憶部f3、次にハイパーリンクテーブルの情報に従ってハイパーテキスト上でハイパーリンクを生成するアドレスハイパーリンク生成部f4を有し、ハイパーテキストa5を出力する。ここでハイパーリンク情報記憶部f3とアドレスハイパーリンク生成部f4は、図2におけるハイパーリンク生成部b3に相当するものの中に構成されている。

【0072】次に動作について説明する。図7のように、DOC1の文字列1からDOC2の文字列2へのハイパーリンクが指定されている場合について説明する。

ただし、文字列のハイパーリンクは、同じファイル内でも可能である。

【0073】この場合、ハイパーリンク情報b2は、図8のハイパーリンク元テーブルg1、ハイパーリンク先テーブルg2のようなハイパーリンク情報となり、DOC1、DOC2等のテキストと共に、ハイパーリンク情報記憶部f3で記憶する。

【0074】図8のハイパーリンクテーブルの情報により、ハイパーリンク元の文字列1を取り出し、同じ番号(N o.)のハイパーリンク先の文字列2へのハイパーリンクを生成する。このID番号は図2におけるハイパーリンク決定部b1により付与されている。

【0075】図8では、ハイパーリンクする文字列を一意に決めるために、ハイパーリンク元文字列とハイパーリンク先文字列が存在する文書(ファイル)名と文字列開始アドレス、文字列終了アドレスを、それぞれハイパーリンク元テーブルg1、ハイパーリンク先テーブルg2に格納している。

【0076】図8のハイパーリンクテーブルでは、格納する順番でハイパーリンク元とハイパーリンク先の対応付けを行なっているので、複数のハイパーリンク元から1つのハイパーリンク先へハイパーリンクを生成する場合は、ハイパーリンク先テーブルに、同じ文字列の情報が複数並ぶことになる。

【0077】以上のように本実施例では、対応させる文字列を順に格納するだけなので、格納が容易である。また、文字列の開始と終了をアドレスで持っているため、文字列の長さに関係なく各データの長さを一定にできる。さらに、文字列が長い場合にもテーブルのデータ量は増えない等の効果がある。

【0078】実施例7. 実施例6におけるハイパーリンクテーブルでは、格納する順番でハイパーリンク元とハイパーリンク先の対応付けを行なっているが、格納順による対応づけでなく、各文字列にIDを付けて、テーブル内に同じ文字列を複数持たないようにして、文字列のIDでハイパーリンクの対応づけをする手法や、ハイパーリンクにIDを付け、ある文字列へハイパーリンクするハイパーリンク元文字列と、そのハイパーリンク先文字列すべてを同じIDとして、そのIDを元にハイパーリンクを生成する手法でも良い。

【0079】以上のように本実施例においては、格納する際にIDを付与したり、格納する際に対応する文字列に同じIDを付与しなければならないが、実施例6で起こるようなデータの重複がないため、データ量が少なくなると共に、IDによってハイパーリンクの対応が明かとなるという効果がある。

【0080】実施例8. 実施例6の図8では、ハイパーリンク元テーブルとハイパーリンク先テーブルを別々に持ったが、その文字列がハイパーリンク元なのかハイパーリンク先なのかの区別がされ、また、IDなどによ

てハイパーリンクの対応関係が取れていれば、1つのテーブルとして持っても良い。本実施例では、1つのテーブルで処理できるという効果がある。

【0081】実施例9. 実施例6のハイパーリンクテーブル図8で、文字列が存在する文書(ファイル)名と文字列開始アドレス、文字列終了アドレスを記憶したのに対して、図9のように、文書(ファイル)名と文字列開始アドレスと、文字列を持つ。この情報によって、文字列開始アドレスから文字列を取り出し、対応する文字列へのハイパーリンクを生成する。同じ字面の文字列で他の文字列にハイパーリンクを生成する場合があるため、文字列だけでなく、文字列開始アドレスが必要である。

【0082】本実施例では以上のように、ハイパーリンクの対象となる文字列のデータをもつため、場合によってはデータ量が増えるが、アドレスだけで表現する場合に比べて、間違いが少なく、または間違いがあってもその検出が容易であるという効果がある。

【0083】実施例10. 実施例9についても、実施例7と同様、文字列のIDや、ハイパーリンクのIDにより、ハイパーリンク元文字列とハイパーリンク先文字列の対応付けを行なっても良い。本実施例においても、実施例7と同様の効果がある。

【0084】実施例11. 実施例9についても、ハイパーリンク元テーブルとハイパーリンク先テーブルを別々に持ったが、実施例8と同様、その文字列がハイパーリンク元なのかハイパーリンク先なのかの区別がされ、また、IDなどによってハイパーリンクの対応関係が取れていれば、1つのテーブルとして持っても良い。本実施例においても、実施例8と同様の効果がある。

【0085】実施例12. 図10は、第12の実施例に係るハイパーテキスト自動生成装置を表す図である。この実施例では、マークアップテキストc4を記憶するマークアップテキスト記憶部h1と、記憶されたマークアップテキストc4のタグの情報に従ってハイパーリンクを生成するタグハイパーリンク生成部h2を有し、ハイパーテキストa5を出力する。ここでマークアップテキスト記憶部h1とタグハイパーリンク生成部h2は、図2のハイパーリンク生成部b3に相当するものの中に構成されている。

【0086】次に動作について説明する。図7のように、DOC1の文字列1からDOC2の文字列2へのハイパーリンクが指定された場合について説明する。ただし、文字列のハイパーリンクは、同じファイル内でも可能である。

【0087】マークアップテキスト記憶部h1に記憶されたマークアップテキストc4は図11のように、ハイパーリンク元文字列とハイパーリンク先文字列の前後にハイパーリンクIDを示すタグを付加したテキストであり、ハイパーリンク元の文字列の前に<start ID=A1>、後ろに、</start>をつけ、ハイパー

10

20

30

40

50

リンク先の文字列の前に<end ID=A1>、後ろに、</end>をつけているが、タグの内容はこれに限らない。文字列の範囲と、ハイパーリンクの元か先の種類、及びハイパーリンクの対応関係がわかればよい。

【0088】次に、タグハイパーリンク生成部h2において、タグの情報を元にハイパーリンクを生成する。IDが同じハイパーリンク元文字列から、ハイパーリンク先文字列へハイパーリンクを生成する。

【0089】本実施例では以上のように、ハイパーリンク情報をテキストに埋め込んでいるため、テーブルとしてではなく、テキストの文字列に情報が付加されているため、人間による変更が容易となる。

【0090】実施例13。図5のマークアップテキストc4はハイパーリンクの情報を含んでいるため、第12の実施例と組み合わせて、第12の実施例の入力である図10のマークアップテキストc4とすることができる。第12の実施例の入力はマークアップテキストだが、この組合せによってベタテキストのハイパーテキスト化が可能となる。

【0091】実施例14。図18は、第14の実施例に係るハイパーテキスト自動生成装置を表す図である。本実施例では、マークアップテキストc4を記憶するマークアップテキスト記憶部h1と、タグテーブルj0より入力された文字列に付けたタグの情報を記憶するタグ情報記憶部j1と、それらの情報に従ってハイパーリンクを生成するタグテーブルハイパーリンク生成部j2を有し、ハイパーテキストa5を出力する。ここで、マークアップテキスト記憶部h1、タグ情報記憶部j1及びタグテーブルハイパーリンク生成部j2は、図2におけるハイパーリンク生成部b3に相当するものの中に構成されている。

【0092】次に動作について説明する。図7のように、DOC1の文字列1からDOC2の文字列2へのハイパーリンクが指定された場合について説明する。ただし、文字列のハイパーリンクは、同じファイル内でも可能である。この場合、マークアップテキストc4は、図11のようになる。実施例12では、マークアップテキスト中のタグの情報を元にハイパーリンクを生成したが、ハイパーリンクを生成する際に、全ファイルからハイパーリンクすべき文字列を探す必要がある。そこで、タグ情報記憶部j1によって、このタグの情報をテーブルとして記憶する。タグテーブルには、図19のように、ハイパーリンクIDと、存在する文書(ファイル)名の関係を格納する。このタグテーブルの情報に従って、タグテーブルハイパーリンク生成部j2によってハイパーリンクを生成する。

【0093】本実施例では以上のように、マークアップとタグ情報記憶等の前処理をしておけば、その後のタグテーブルハイパーリンク生成部におけるハイパーリンク

生成が高速に行える。

【0094】実施例15。図5のマークアップテキストc4はハイパーリンクの情報を含んでいるため、第14の実施例と組み合わせて、第14の実施例の入力である図18のマークアップテキストc4とすることができる。第14の実施例の入力はマークアップテキストだが、この発明の組合せによってベタテキストのハイパーテキスト化が可能となる。

【0095】実施例16。実施例14におけるタグテーブルでは、ハイパーリンクIDと、存在する文書(ファイル)名の関係を格納しているが、文書量や、ハイパーリンクを生成するマシンの状況に応じて、タグテーブルは、図20のように、ハイパーリンクID、文字列、存在する文書(ファイル)名、位置情報、など必要事項を格納しても良い。本実施例では、データ量は多くなるが、間違い検出が容易となる。

【0096】実施例17。第4の実施例のテキストa1は、テキストではなくテキスト集合やマークアップテキストやマークアップテキスト集合でも良い。すなわち、第3の実施例に係るカード集合c7でも良い。従って、第3の実施例と第4の実施例を組み合わせることによって図3のテキストa1から図4のキーワード集合a3を出力することができる。

【0097】実施例18。実施例3はハイパーリンク元とハイパーリンク先を定めたタイトルハイパーリンク情報c8を出力するので、実施例6と組み合わせることによって、上位カードに残したタイトル文字列と一段下位のカードに残したタイトル文字列の間を自動ハイパーリンクすることができる。

【0098】実施例19。実施例3におけるカード集合c7はマークアップテキストc4の形式なので、第12の実施例と組み合わせることによって、上位カードに残したタイトル文字列と一段下位のカードに残したタイトル文字列の間を自動ハイパーリンクすることができる。このような組み合わせを行う場合は、タイトルハイパーリンク情報c8を出力する必要はない。

【0099】実施例20。実施例3は実施例14と組み合わせることによって、上位カードに残したタイトル文字列と一段下位のカードに残したタイトル文字列の間を自動ハイパーリンクすることができる。

【0100】実施例21。実施例4は、キーワード集合a3を出力するが、これを実施例5におけるキーワード集合a3と同じ形式にしておくことで実施例4と実施例5に係るハイパーテキスト自動生成装置とを組み合わせることことができる。これによって、キーワード集合a3を手作業作成することなく、テキストを自動ハイパーリンクすることが可能となる。

【0101】なお、実施例4と実施例5における図12、図14、図16とを組み合わせる場合はキーワード抽出部d2においては出現頻度が1回の文字列はハイパ

ーリンク可能文字列候補d 3から除いて出力することができる。同じ文字列をハイパーリンクする場合は、同じ文字列がハイパーリンク元とハイパーリンク先の最低2回出現することが必要であるからである。

【0102】実施例22、図21は、実施例3から実施例14に係る装置を組み合わせたハイパーテキスト自動生成装置を表す図である。実施例3、実施例4、実施例5及び実施例6から実施例14までの実施例の1つを組み合わせることによって、テキストからハイパーテキストを自動生成する自動化の度合を最も高めることができる。

【0103】実施例23、図22は第23の実施例に係るハイパーテキスト自動生成装置を表す図である。図22においてb1はハイパーリンク決定部、m3はある戦略1にそったハイパーリンクを決定する戦略1の戦略別ハイパーリンク決定部、m4はある戦略2にそったハイパーリンクを決定する戦略2の戦略別ハイパーリンク決定部、m2は戦略を表す戦略文字列m1を入力し、戦略1、戦略2、・・・の中から適切な戦略を選択する戦略選択部である。ここで戦略とは、リンクを生成するときの方針を表現したものである。

【0104】次に動作について説明する。戦略を表す文字列m1がハイパーリンク決定部b1に入力されると、戦略選択部m2ではハイパーリンク決定部b1内にある複数の戦略別ハイパーリンク決定部m3、m4、・・・m5の中から、適切なものを選択する。選択された戦略別ハイパーリンク決定部は、テキストa1とハイパーリンク可能文字列候補d3が入力されると、各ハイパーリンク可能文字列候補について、そのハイパーリンク元文字列とハイパーリンク先文字列が前記実施例5の手法に従って決定され、マークアップテキストc4およびハイパーリンク情報b2として出力される。マークアップテキストc4またはハイパーリンク情報b2は、その片方だけでもハイパーリンク元およびハイパーリンク先の決定情報を含むので、ハイパーリンク決定部b1の出力としては、マークアップテキストc4またはハイパーリンク情報b2のいずれか片方でも良い。

【0105】図23は、第23の実施例に係るハイパーリンクを示す概念図である。DOC1という文書内の文字列「A型部品」をハイパーリンク元文字列とするハイパーリンクを生成する場合について説明する。ここで戦略1、戦略2は、いくつか用意されている戦略のうちのそれぞれ1つである。ここで戦略とは、前記したようにリンクを生成するときの方針を表現したものである。例えば、戦略1は「ハイパーリンク先文字列は、次に同じ文字列が現れる文書のタイトルとする」であり、同じく、戦略2は「ハイパーリンク先文字列は、同じ文字列が多数現れる文書の最初の文字列とする」である。戦略の名前は、戦略選択部m2の中で対応がとれていれば、戦略1、戦略2というものでなくても、どのような文字列でもよい。例えば戦略の内容を想起できるように、この例

の戦略1は「チェーン」、戦略2は「詳細説明」というようにつけておいてもよい。

【0106】図23において、戦略文字列として「戦略1」が与えられたとするとハイパーリンク元文字列に対して、ハイパーリンク先は、DOC2のタイトル文字列となる。また「戦略2」が与えられたとすると、ハイパーリンク元文字列に対して、ハイパーリンク先は、DOC3の最初の「A型部品」という文字列となる。

【0107】もちろん戦略はハイパーリンク元文字列の決定にも適用できる。例えば、戦略3は「ハイパーリンク元文字列は、文書内の全ての同じ文字列とする」のことである。図23において、戦略文字列として「戦略3」が与えられたとするとDOC1内のすべての文字列「A型部品」を、ハイパーリンク元文字列とする。

【0108】本実施例では以上のように、用途により異なる戦略を与えることにより、適切なハイパーリンクを選ぶことが可能となるので、同じテキストについてハイパーリンクの自動生成を行っても、必ず同じハイパーテキストを得るのではなく、用途に応じたハイパーテキストを得ることができる。

【0109】実施例24、図24は第24の実施例に係るハイパーテキスト自動生成装置を表す図である。図においてb1はハイパーリンク決定部、n1はハイパーリンクの挙動を表す属性を決定する属性決定部である。ここで挙動とは、ハイパーリンクが張られているハイパーリンク元文字列をクリックしたときに、ハイパーリンク先をどのように表示し、次にどうなるか、という一連の動作のことである。例えば、ハイパーリンク元文字列をクリック後、ハイパーリンク先を表示して、ハイパーリンク元を消去する、というものである。あるいは、ハイパーリンク元文字列をクリック後、ハイパーリンク元文字列に重ねてハイパーリンク先を表示して、もう1度クリックすると、ハイパーリンク先を消去して、ハイパーリンク元文字列を表示する、というものである。

【0110】次に動作について説明する。まず、各ハイパーリンク可能文字列候補d3について、そのハイパーリンク元文字列とハイパーリンク先文字列が前記実施例5の手法に従って決定される。ここで属性決定部n1では、決定されたハイパーリンク情報をもとに、どのような出現パターンの場所にリンク作成されたかを判断し、属性を決定し、ハイパーリンク情報に付与して、マークアップテキストn2およびハイパーリンク情報n3として出力される。マークアップテキストn2またはハイパーリンク情報n3は、その片方だけでもハイパーリンク元およびハイパーリンク先の決定情報を含むので、ハイパーリンク決定部b1の出力としては、マークアップテキストn2またはハイパーリンク情報n3のいずれか片方でも良い。

【0111】図25は、第24の実施例に係るハイパーリンクを示す概念図である。DOC1内の文字列「A型

部品」からDOC 2内のタイトル文字列にリンク生成を行なうようなハイパーリンク情報を決定したとする。この時、文書を代表するタイトルヘリンクが作成されと

いうことで、「参照」という属性を与える。すると、「参照」の場合は次の文書が読みたいという挙動をとらせることとし、DOC 1が表示されている画面で、文字列「A型部品」をクリックすると、新たな画面では、DOC 1は消えDOC 2が表示される。同様に、次の画面ではDOC 2が消えDOC 3が表示される。

【0112】DOC 1内の文字列「A型部品」からDOC 3内の文字列「A型部品」にリンク生成を行なうようなハイパーリンク情報を決定したとする。この時、1文書内に多数現れる文字列に対してリンクが作成されたという

ことで、「注釈」という属性を与える。すると、「注釈」の場合は一時的にリンク先の文書を見たいという挙動をとらせることとし、DOC 1が表示されている画面で文字列「A型部品」をクリックすると、新たな画面では、DOC 1は残りその上に重ねるようにDOC 3が表示される。次の画面では、DOC 3のみが消え、再びDOC 1が表示される。

【0113】本実施例では以上のように、ハイパーリンク情報の属性を自動的に決定できるようにしたので、その属性により表示形態を変えるハイパーテキストも自動的に生成することができる。

【0114】

【発明の効果】第1の発明は、入力したテキストを所定の単位に分解し所定の文字列を抽出する文字列抽出決定部と、抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と一意に関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成するハイパーリンク部とを備えているので、索引文字列文書ではなく、ベタのテキストを入力してもハイパーテキストの自動生成が可能となる。また、テキスト対テキストや文字列対テキストの自動ハイパーリンクではなく、文字列から文字列への自動ハイパーリンクを可能にし、文字列単位の詳細なハイパーリンクを可能にする。

【0115】第2の発明は、ハイパーリンク部を、ハイパーリンク元文字列とハイパーリンク先文字列を決定し両者を関連付けるハイパーリンク情報を出力するハイパーリンク決定部と、ハイパーリンク情報により前記ハイパーリンクを生成するハイパーリンク生成部とを分離して備えているので、ハイパーリンクの決定と、ハイパーリンクの生成とを別のマシンで行ったり、それぞれを別々のアプリケーションで実現可能となる。

【0116】第3の発明は、テキストを入力しハイパーリンク可能な文字列候補を抽出する文字列抽出部と、文字列候補を所定の順序に並べユーザの指定によりハイパーリンクする文字列を決定する文字列決定部とを備えているので、ハイパーリンクする文字列候補抽出の自動化

が可能となり、実際にハイパーリンクする文字列の決定はユーザが行うことで、有効なハイパーリンクだけを生成することができる。

【0117】第4の発明は、テキスト中の抽出された文字列をハイパーリンク元とし、そのテキストとは別のテキストのタイトル文字列に出現するハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定するので、ハイパーリンク元文字列に関する詳しい情報が記述されている可能性の高いテキストへハイパーリンクすることが可能となる。

【0118】第5の発明は、テキスト中の抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列が最も多く出現するハイパーリンク元のテキストとは別のテキストのタイトル文字列をハイパーリンク先として決定するので、ハイパーリンク元文字列に関する詳しい情報が記述されている可能性の高いテキストへハイパーリンクすることが可能となる。

【0119】第6の発明は、テキスト中の抽出された文字列をハイパーリンク元とし、そのハイパーリンク元の文字列と同一文字列または同義の文字列をハイパーリンク先として決定するので、同一または同義の文字列に関する記述を順次閲覧することが可能となる。

【0120】第7の発明は、テキストのハイパーリンク元文字列とハイパーリンク先文字列をテキストのアドレスにより一意に決める情報及びハイパーリンク元文字列とハイパーリンク先文字列の対応関係の情報とを含んだハイパーリンク情報を記憶するハイパーリンク情報記憶部と、そのハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成するアドレスハイパーリンク生成部とを備えているので、ハイパーリンクの決定内容を変更する際に、最初からやり直さなくてもハイパーリンク情報記憶部の情報を変更すれば良い。

【0121】第8の発明は、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力し、マークアップテキストを記憶するマークアップテキスト記憶部と、そのマークアップテキスト記憶部に記憶されたマークアップテキストのタグ情報によりハイパーリンクを生成しており、テキストの文字列に付加されたハイパーリンク情報を使用しているため、ハイパーリンクの決定内容を変更する時にユーザによる変更が容易となる。

【0122】第9の発明は、ハイパーリンク元文字列とハイパーリンク先文字列のハイパーリンクの対応関係を示すタグ情報を付加したマークアップテキストを出力し、マークアップテキストを記憶するマークアップテキスト記憶部と、タグ情報をテーブルとして記憶するタグ情報記憶部と、そのタグ情報記憶部に記憶された前記タ

グ情報によりハイパーリンクを生成するので、ハイパーリンクを生成する際に全テキストからハイパーリンクすべき文字列を探す必要がなく、ハイパーリンクの生成が高速に行える。

【0123】第10の発明は、入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力する文書構造解析部と、マークアップテキストを入力し指定された所定の分割単位にマークアップテキストを分割してタグ情報を付加した部分マークアップテキスト群であるカード集合を出力するとともに、カード集合におけるハイパーリンク元文字列とハイパーリンク先文字列を定めたハイパーリンク情報とを出力するカード分割部と、ハイパーリンク情報により、カード集合のハイパーリンクを生成しハイパーテキストを出力するハイパーリンク生成部とを備えているので、テキストを適切な表示単位であるカードに自動的に分割することが可能となる。

【0124】第11の発明は、ハイパーリンク情報を記憶するハイパーリンク情報記憶部を備え、そのハイパーリンク情報記憶部に記憶されたハイパーリンク情報によりハイパーリンクを生成するので、ハイパーリンクの決定内容を変更する際に、最初からやり直さなくてもハイパーリンク情報記憶部の情報を変更すれば良い。

【0125】第12の発明は、入力したテキストの論理構造を所定の文書型定義にもとづき解析し、その論理構造のタグ情報を付加したマークアップテキストを出力する文書構造解析部と、マークアップテキストを入力し指定された所定の分割単位にマークアップテキストを分割して前記タグ情報を付加した部分マークアップテキスト群であるカード集合を出力するカード分割部と、カード集合を入力しタグ情報が付加された文字列を抽出する文字列抽出決定部と、抽出された所定の文字列をハイパーリンク元とし、そのハイパーリンク元の文字列とタグ情報により関連付けられた文字列をハイパーリンク先として決定し、ハイパーリンクを生成しハイパーテキストを出力するハイパーリンク部とを備えているので、テキストを適切な表示単位であるカードに自動的に分割することが可能となる。

【0126】第13の発明は、カード集合を記憶し、その記憶部に記憶されているカード集合のタグ情報によりハイパーリンクを生成しており、テキストの文字列に付加されたハイパーリンク情報を使用しているため、ハイパーリンクの決定内容を変更する際に、ユーザによる変更が容易となる。

【0127】第14の発明は、カード集合を記憶部に記憶し、カード集合のタグ情報をテーブルとしてタグ情報記憶部に記憶し、そのタグ情報記憶部に記憶されたタグ情報によりハイパーリンクを生成するので、ハイパーリンクを生成する際に、全テキストからハイパーリンクす

べき文字列を探す必要がなく、ハイパーリンクの生成が高速に行える。

【0128】第15の発明は、リンクを生成するときの方針を示した戦略文字列にもとづき、適切なハイパーリンクを決定するように構成したので、同じテキストについてのハイパーリンクの自動生成を行なっても、必ず同じハイパーテキストを得るのではなく、用途に応じたハイパーテキストを得ることが可能となる。

【0129】第16の発明は、ハイパーリンク先の文字列にもとづきハイパーリンク先のテキストの表示方法の属性を決定するように構成したので、その属性により表示形態を変えるハイパーテキストも自動的に生成することが可能となる。

【図面の簡単な説明】

【図1】 この発明の第1の実施例のハイパーテキスト自動生成装置である。

【図2】 この発明の第2の実施例のハイパーテキスト自動生成装置である。

【図3】 この発明の第3の実施例のハイパーテキスト自動生成装置である。

【図4】 この発明の第4の実施例のハイパーテキスト自動生成装置である。

【図5】 この発明の第5の実施例のハイパーテキスト自動生成装置である。

【図6】 この発明の第6の実施例のハイパーテキスト自動生成装置である。

【図7】 この発明の文字列のハイパーリンクを表す図である。

【図8】 この発明の第6の実施例のハイパーリンクテーブル1の内容を表す図である。

【図9】 この発明の第9の実施例のハイパーリンクテーブル2の内容を表す図である。

【図10】 この発明の第12の実施例のハイパーテキスト自動生成装置である。

【図11】 この発明の第12の実施例の文書にタグを付加した状態を表す図である。

【図12】 この発明の第5の実施例におけるハイパーリンクを示す概念図である。

【図13】 この発明の第5の実施例におけるハイパーリンクを示す概念図である。

【図14】 この発明の第5の実施例におけるハイパーリンクを示す概念図である。

【図15】 この発明の第5の実施例におけるハイパーリンクを示す概念図である。

【図16】 この発明の第5の実施例におけるハイパーリンクを示す概念図である。

【図17】 この発明の第5の実施例におけるハイパーリンクを示す概念図である。

【図18】 この発明の第14の実施例のハイパーテキスト自動生成装置である。

25

【図19】 この発明の第14の実施例のタグテーブル1の内容を表す図である。

【図20】 この発明の第16の実施例のタグテーブル2の内容を表す図である。

【図21】 この発明の第22の実施例のハイパーテキスト自動生成装置である。

【図22】 この発明の第23の実施例のハイパーリンク自動生成装置である。

【図23】 この発明の第23の実施例のハイパーリンクを示す概念図である。

【図24】 この発明の第24の実施例のハイパーリンク自動生成装置である。

【図25】 この発明の第24の実施例の動作を示す図である。

【図26】 従来のハイパーテキストの概念を説明するための図である。

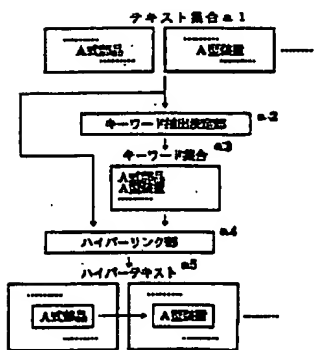
【符号の説明】

a1 テキスト集合、テキスト、a2 キーワード抽出決定部、a3 キーワード集合、a4 ハイパーリンク部、a5 ハイパーテキスト、b1 ハイパーリンク決定部、b2 ハイパーリンク情報、b3 ハイパーリンク生成部、c2 文書型定義、c3 文書構造解析部、c4 マークアップテキスト、c5 分割単位指定、c6

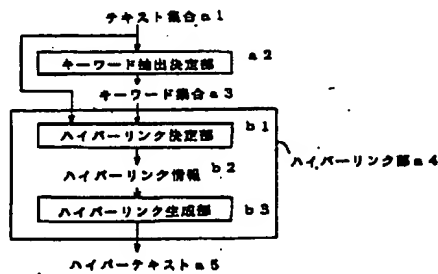
26

カード分割部、c7 カード集合、c8 タイトルハイパーリンク情報、d2 キーワード抽出部、d3 ハイパーリンク可能文字列候補、d4 キーワード決定部、f3 ハイパーリンク情報記憶部、f4 アドレスハイパーリンク生成部、g1 ハイパーリンクテーブル1のハイパーリンク元テーブル、g2 ハイパーリンクテーブル1のハイパーリンク先テーブル、g3 ハイパーリンクテーブル2のハイパーリンク元テーブル、g4 ハイパーリンクテーブル2のハイパーリンク先テーブル、h1 マークアップテキスト記憶部、h2 タグハイパーリンク生成部、i1 タグ情報を付加したハイパーリンク元文字列、i2 タグ情報を付加したハイパーリンク先文字列、j0 タグテーブル、j1 タグ情報記憶部、j2 タグテーブルハイパーリンク生成部、k1 タグテーブル1のハイパーリンク元テーブル、k2 タグテーブル1のハイパーリンク先テーブル、k3 タグテーブル2のハイパーリンク元テーブル、k4 タグテーブル2のハイパーリンク先テーブル、m1 戦略文字列、m2 戦略選択部、m3 戦略1のハイパーリンク決定部、m4 戦略2のハイパーリンク決定部、n1 属性決定部、n2 マークアップテキスト、n3 ハイパーリンク情報。

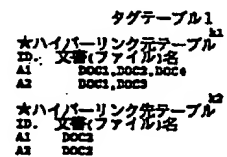
【図1】



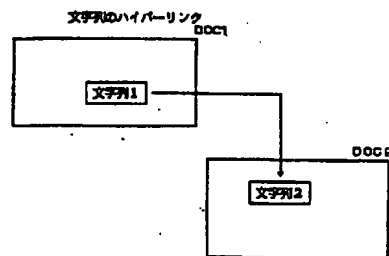
【図2】



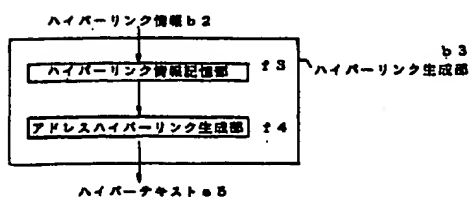
【図19】



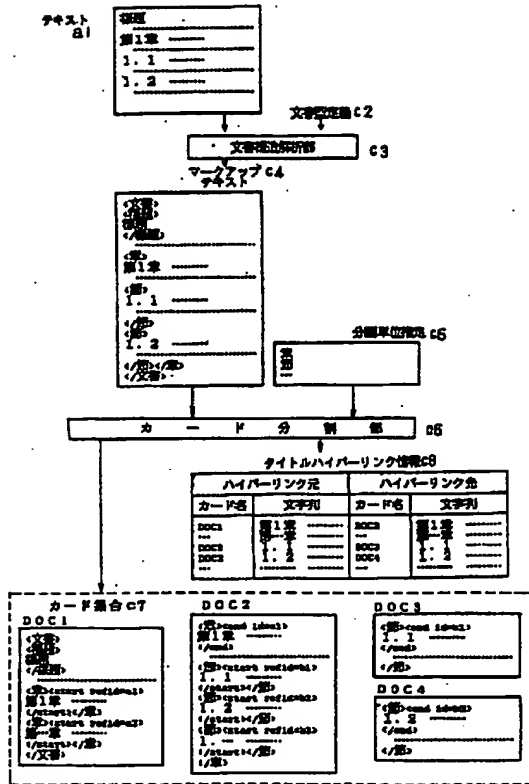
【図7】



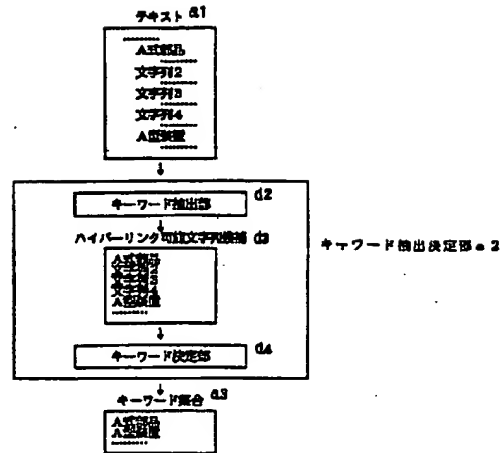
【図6】



【図3】



【図4】



【図8】

ハイパーリンクテーブル1

NO.	文書(ファイル)名	文字列開始アドレス	文字列終了アドレス
1.	DOC1	48000	50000
2.	DOC2	50000	52000

★ハイパーリンク先テーブル

NO.	文書(ファイル)名	文字列開始アドレス	文字列終了アドレス
1.	DOC1	50000	52000
2.	DOC2	52000	54000

【図9】

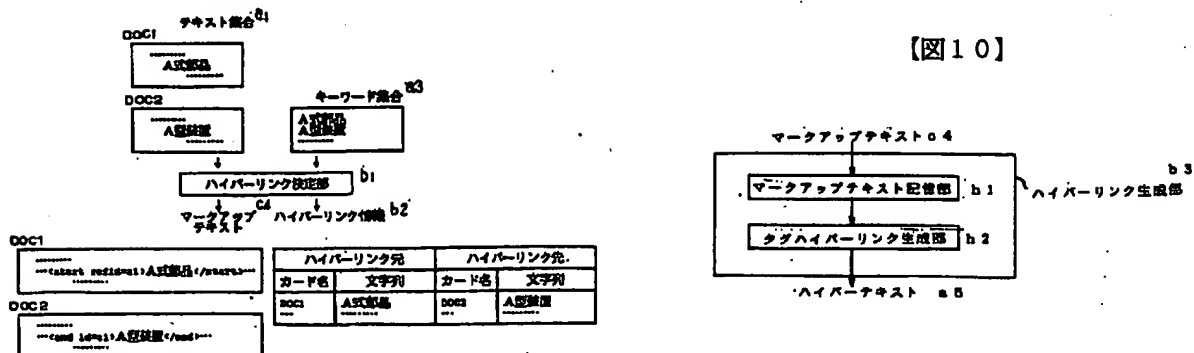
ハイパーリンクテーブル2

NO.	文書(ファイル)名	文字列開始アドレス	文字列終了アドレス
1.	DOC1	48000	50000
2.	DOC2	50000	52000

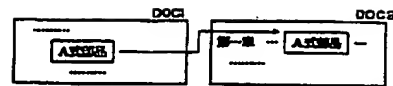
★ハイパーリンク先テーブル

NO.	文書(ファイル)名	文字列開始アドレス	文字列終了アドレス
1.	DOC1	50000	52000
2.	DOC2	52000	54000

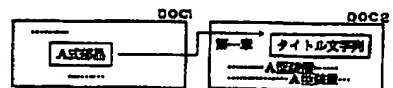
【図10】



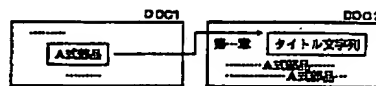
【圖 12】



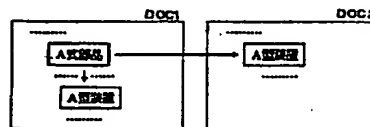
【図 15】



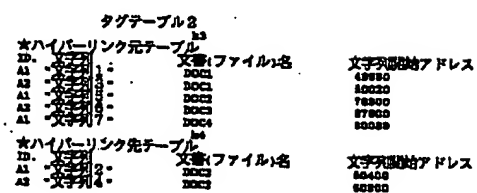
【图 14】



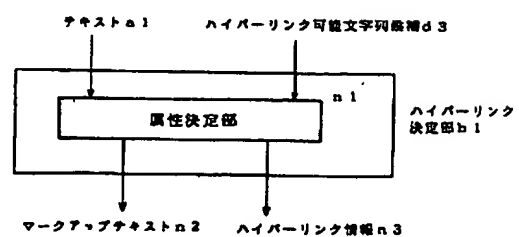
【図 17】



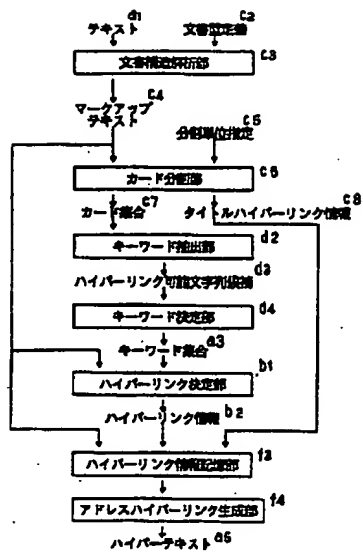
【図20】



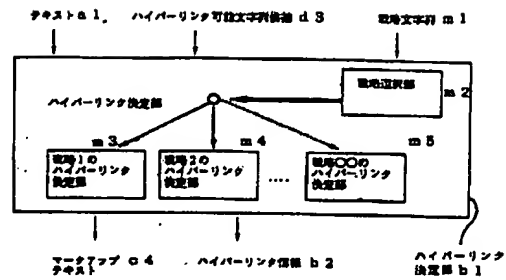
【图 24】



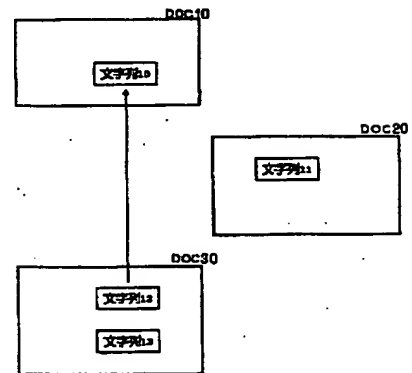
【図21】



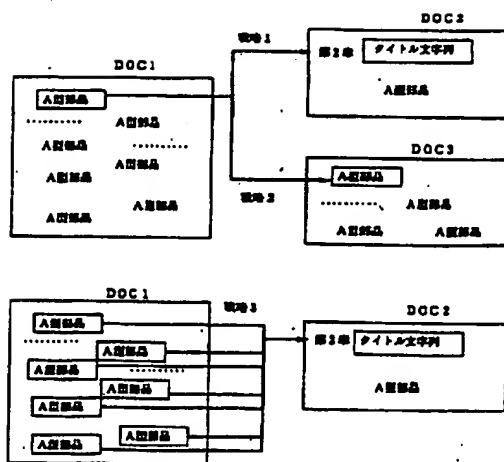
【図22】



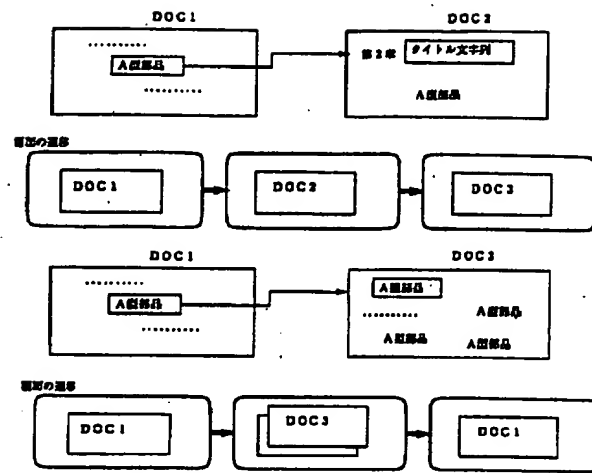
【図26】



【図23】



【図25】



フロントページの続き

(51)Int.Cl.⁶
G 0 6 F 17/30

識別記号

庁内整理番号

F I

技術表示箇所

9288-5L

G 0 6 F 15/20

5 7 0 R

9194-5L

15/419

3 1 0

(72)発明者 田中 聡

鎌倉市大船五丁目1番1号 三菱電機株式
会社情報システム研究所内